An observing system simulation experiment to evaluate the scientific merit of wind and ozone measurements from the future SWIFT instrument

By W. A. LAHOZ[1]*, R. BRUGGE[1], D. R. JACKSON[2], S. MIGLIORINI[1], R. SWINBANK[2], D. LARY[3] and A. LEE[4]

[1]*Data Assimilation Research Centre, University of Reading, UK*
[2]*Met Office, Exeter, UK*
[3]*Global Modeling and Assimilation Office, NASA GFSC, Greenbelt, MD, USA*
[4]*Singapore University, Singapore*

SUMMARY

An observing system simulation experiment was performed to assess the impact and scientific merit of SWIFT stratospheric wind and ozone observations. The SWIFT instrument is being considered for launch later this decade, and is expected to provide unprecedented global information on key aspects of the stratosphere, including tropical winds, ozone fluxes and wintertime variability. It was found that SWIFT wind observations will have a significant impact on analyses in the tropical stratosphere (except the lowermost levels), and could have a significant impact in the extratropics when the SWIFT observations are available and the flow regime is changing relatively fast. Results indicate that SWIFT ozone observations will have a significant impact when the vertical gradient of ozone is relatively high. The experiments indicate that SWIFT wind observations would improve the analysis of both tropical wind and wintertime variability. The results of this study strongly indicate a beneficial impact from the proposed SWIFT instrument.

KEYWORDS: Data assimilation Satellite OSSE Remote sensing

1. INTRODUCTION

The current operational meteorological observing system includes a wide variety of measurements, from both *in situ* and remote-sensing instruments. The network of operational radiosondes forms the backbone of the conventional observing system, providing regular profiles of wind and temperature through the troposphere and lowermost stratosphere. The conventional observing system is complemented by an increasingly important set of remote-sensing observations, mainly from satellite instruments. These include atmospheric motion vectors and nadir soundings. The atmospheric motion vectors are primarily available for the troposphere, while the operational nadir soundings provide temperature-profile information through both the troposphere and stratosphere. Thus, there are essentially no operational observations of wind vectors above the level reached by the radiosondes.

To fill this gap in the observing system, the SWIFT (Stratospheric Wind Interferometer For Transport studies) instrument has been proposed (Shepherd *et al.* 1998), for possible flight later this decade. SWIFT is a limb-sounding instrument, which would measure the line-of-sight velocity using the Doppler shift of a particular ozone spectral line (1133.4335 cm^{-1}; see, for example, Dobbie *et al.* (1996)). A pair of measurements of the same location, viewed from two perpendicular directions, are made in quick succession; this enables the horizontal vector wind to be derived. Since SWIFT uses an ozone emission line, ozone measurements may be derived at the same time as the wind measurements; it is thus a particularly relevant instrument for studies of ozone transport.

A previous limb-viewing stratospheric wind instrument, HRDI (High Resolution Doppler Imager; Hays *et al.* (1993)), was carried on the NASA Upper Atmosphere Research Satellite (UARS). In a study of the assimilation of HRDI measurements,

---

Boorman *et al.* (2000) found that the impact of the HRDI data was marginal, since the observation errors were large compared with the estimated forecast wind errors. (The stratospheric winds can be derived, to a reasonable approximation, from the satellite temperature soundings, by the assimilation model.) However, the proposed SWIFT instrument will produce more accurate measurements.

As discussed by Randel *et al.* (2002), we do not currently have a good estimate of the state of the tropical stratosphere. In particular, the quasi-biennial oscillation (QBO) variability is underestimated, and 'balanced winds' in the tropics are problematic for estimating the variability in the QBO. SWIFT is expected to improve the quality of stratospheric analyses in the tropics.

It is now clear that human activity is directly responsible for the destruction of the ozone layer through the anthropogenic release of ozone-destroying chemicals (WMO 1999). To understand the mechanisms responsible for ozone change, it is important to distinguish between chemical and dynamical effects. By providing simultaneous measurements of winds and ozone, SWIFT is expected to help improve understanding of the link between dynamics and chemistry occurring through transport of chemical species, including ozone. By improving the quality of stratospheric analyses (especially in the tropics) SWIFT is also expected to help support research on climate change by providing quality-controlled analyses to evaluate model results and initialize model runs.

The aim of this study is to assess whether the SWIFT instrument can have a significant beneficial impact. The present paper focuses on the scientific results of the study. Lahoz *et al.* (2003) have provided further details.

A standard way to assess a proposed addition to the observing system is to carry out an observing system simulation experiment (OSSE; e.g. Atlas 1997). The first component of an OSSE is a 'nature run' (or reference atmosphere); this is normally produced by integrating a general-circulation model (GCM). In this paper the nature run comes from meteorological analyses. Then, a complete set of observations is simulated from the nature run. These observations are a complete reproduction of the operational observing network (or the expected configuration of the network at some future time). In addition, the measurements from the proposed new observation type are also simulated. Two assimilation experiments are run, one with a data assimilation system using all the simulated operational observations, and the second using the new observations in addition to all the operational observations. Both assimilation experiments are then compared with one another, and with the reference atmosphere, to assess the impact of the new observation type.

Setting up a system for carrying out a complete simulation of all the different observation types is a major undertaking—comparable to writing the assimilation system itself. The performance and evaluation of the assimilation experiments is also a time-consuming exercise. Nevertheless, it is often a worthwhile undertaking, especially considering the possible benefit in the evaluation of very expensive observing systems.

The study described in this paper was funded by the European Space Agency with a limited budget for the evaluation of the likely benefit of SWIFT. While our team had access to a good operational stratospheric data assimilation system from the Met Office (Swinbank and O'Neill 1994; Swinbank *et al.* 2002), we did not have the basic infrastructure to carry out OSSEs with the full simulation of the observing system. Instead, we have used an approach we refer to as a 'reduced OSSE', where we simulate a reduced version of the observing system. For example, instead of simulating satellite radiance observations, we simulate profiles of retrieved profiles of temperatures, at a reduced horizontal resolution (consistent with our model resolution).

While this approach does not follow the best practice for OSSEs, we judge that this should be perfectly satisfactory for the SWIFT data, since they are a completely different observation type from the other observations available in the operational observing system. If one were evaluating a new observation type that was essentially an incremental improvement on previous observation types, one would need to undertake a full sophisticated (and expensive) OSSE to obtain credible results.

In the next section we describe in detail the experimental set-up, including the production of the reference atmosphere, simulation of the observations, the assimilation system and the evaluation of the OSSE. We then examine the impact of the simulated SWIFT observations in the assimilation experiments. Finally, we assess the likely scientific merits of the proposed SWIFT instrument.

2.   EXPERIMENTAL SET-UP

(a)   *Reference atmosphere*

The OSSE was created and run for two month-long periods in 2000 (January and April). A primary reason for selecting these periods is that they are scientifically interesting, and include a sufficient variety of conditions to address the scientific issues associated with the SWIFT instrument. Measurements from the THESEO* field campaign (EC 2000) made during the 1999–2000 northern hemisphere (NH) winter showed the largest springtime ozone losses hitherto recorded in the NH polar vortex, with local losses approaching 70% at 20 km when compared with early winter values. January temperatures in the polar lower stratosphere reached new historical lows, which led to increased polar stratospheric cloud (PSC) formation and significant levels of denitrification and dehydration. After the final warming in mid-March, ozone-depleted polar air mixed with mid-latitude air, and this process continued throughout April.

The choice of the 2000 NH winter and spring also allowed the study to make use of the 60-level meteorological analyses from the European Centre for Medium-Range Weather Forecasts (ECMWF) assimilation system, an important consideration in order to minimize the problem of 'incest', where the same model is used to generate the reference atmosphere and perform the assimilation. These data are available globally every six hours, with a vertical range from the surface to about 65 km and a vertical resolution in the stratosphere of 1.5 km. The spherical harmonics were truncated using a triangular filter at T42 resolution and then these spectral coefficients were transformed onto a Gaussian grid at a horizontal resolution in physical space of approximately $2.8° \times 2.8°$.

In order to generate all the observations to be included in the OSSE, the reference atmosphere must include horizontal wind, temperature, specific humidity and ozone. The first three of these are available from the ECMWF analyses. The ozone fields had to be simulated using the SLIMCAT chemical transport model (Chipperfield 1999). Here, the model circulation is driven by ECMWF analyses, and the model transport code is coupled to a detailed stratospheric chemistry scheme, which includes a treatment of heterogeneous reactions on both PSCs and sulphate aerosols.

The distributions of chemical species used in the simulation were initialized on 9 December 1999 from a low-resolution multi-annual simulation (Chipperfield 1999). These simulations were run at T31 horizontal resolution (approximately $3.75° \times 3.75°$) and at 18 isentropic levels between 335 and 2700 K (approximately 10 to 60 km) at

* THird European Stratospheric Experiment on Ozone.

TABLE 1.   SUMMARY OF NON-SWIFT SATELLITE-BASED SIMULATED OBSERVATIONS AND THEIR CHAR-
ACTERISTICS DURING ANY SIX-HOUR PERIOD

Instrument	Observation	Average number of locations	Horizontal coverage	Horizontal resolution	Vertical resolution (km)	Vertical data range (km)
IASI	T	432	Two 45° longitude bands (one satellite)	5° × 5°	1	0–20
	RH				1	0–10
	O_3				4–17	15–40
AMSU-A	T	1296	Four 45° longitude bands (two satellites)	5° × 5°	8	0–45
GRAS	T	125	Global	Nominally 5° × 5°, but random spread across globe	1.5 3	0–5 5–50
GOME-2	O_3	238	Two 45° longitude bands (one satellite)	Nominally 5° × 5° but solar zenith angle <75°	5–10	0–50
SEVIRI	**u**	1800	60°N–60°S	5° × 5°	Single level	Cloud top only
ASCAT	**u**	1305 (max)	60°N–60°S, where cloud-free	Irregular, over ocean only	Single level	At surface

Observations provided at 00, 06, 12 and 18 UTC.
T = Temperature, RH = Relative humidity, O_3 = ozone, **u** = wind direction and speed.

intervals of between 1.5 km and 2.0 km in the lower stratosphere. The ozone fields were stored every six hours on these isentropic levels.

An investigation of the reference-atmosphere fields (not shown) showed them to be in reasonable agreement with climatological data and with independent observations of the period in question.

(*b*)   *Simulated observations*

For this study, we simulated both SWIFT observations and a set of standard operational observations. These sets of observations were used in parallel assimilation experiments: SW (SWIFT) included both the operational and SWIFT observations, and NS (no SWIFT) only included operational data. These data are characterized in Tables 1–3. General details regarding the simulation of the observations follow. Further details on the simulation of the SWIFT data appear in section 2(c).

The simulated data were created by interpolating horizontally between the four closest grid points of the reference atmosphere onto the observation location, and then vertically onto the observation measurement levels. The interpolation is bi-linear in the horizontal and linear in log-pressure in the vertical.

Prior to this interpolation, gridded fields of relative humidity (RH) and potential temperature also need to be calculated. The RH is needed to calculate the relevant observations in Tables 1 and 2, while the potential temperature is important in the calculation of the ozone observations. This is because the gridded ozone data are only available between the 335 and 2700 K isentropic surfaces. At each observation location

TABLE 2. SUMMARY OF SIMULATED NON-SATELLITE OPERATIONAL OBSERVATIONS AND THEIR CHARACTERISTICS DURING ANY SIX-HOUR PERIOD

Instrument	Average number of locations	Horizontal coverage	Vertical resolution	Vertical data range	Temporal coverage
Surface	2522 buoys 1340 ships 6101 land	Global (as current observing system)	Single level	Surface only	Continuous, but mainly at 00, 06, 12 and 18 UTC for ship and land stations
Aircraft	10000 AMDARs 4120 AIREPs	Global (as current flight corridors)	Single level	Flight level	Continuous
Radiosonde	306 TEMPs 236 PILOTs	Global (as current observing system)	Irregular	Mostly troposphere	Mainly at 00 and 12 UTC but a few at 06 and 18 UTC

Data on pressure, relative humidity, temperature, wind speed and direction provided.

the potential temperature was first calculated, and from this the points at which no valid ozone data were available were identified and flagged.

In any six-hour period, the number of profiles containing satellite data far exceeds the number of grid points in the assimilation system; here the model used has 96 points in longitude and 73 in latitude. As a consequence, in any operational system the satellite data are routinely thinned. In the operational Met Office stratospheric assimilation system, these data are thinned to $2° \times 2°$ bins. Thus, in order to avoid simulating a wealth of satellite data that would not be used in the assimilation, all the NS satellite data were pre-thinned to one profile per instrument (IASI, AMSU-A, GRAS, GOME-2) per $5° \times 5°$ bin. Note that the thinned data from these four instruments, when grouped together, have a coverage of around $2.5° \times 2.5°$, which is close to the density of data used in the operational Met Office system.

Due to an oversight, a value of 8 km was used for the vertical resolution of the AMSU-A retrievals, when the correct value is about 5 km. This will tend to underestimate the impact of AMSU-A data, most particularly in regions where the temperature gradient is nonlinear and/or non-monotonic (e.g. the tropopause and the stratopause). However, the vertical observation-error correlations in the AMSU-A retrievals were disregarded, and this more than likely offset the negative impact of using too coarse a vertical resolution.

The errors were calculated by multiplying the published standard deviations (STDs) for each observation type or instrument by a randomly generated number. These numbers have an output set that has a Gaussian distribution with zero mean and unit variance. Each individual error was then added to the 'base value' of each observation, which was derived from the reference atmosphere. For ozone and RH, the addition of the observation errors was constrained to ensure no negative-valued ozone observations or supersaturated RH observations appeared. The simulated observations are assumed to be unbiased.

The observation errors for the aircraft, sonde and surface data include representativeness error. The size of this error has been determined through experience of using these data in the Met Office operational numerical weather prediction (NWP) system. However, the other instruments are new or planned and, given that the resolutions of

TABLE 3. SUMMARY OF SIMULATED SWIFT OBSERVATIONS AND THEIR CHARACTERISTICS DURING ANY SIX-HOUR PERIOD

Observation	Vertical data range (km)	Vertical resolution (km)	Number of locations	Horizontal coverage
O_3	15–45	2	360	87°N to 53°S or 53°N to 87°S
u	15–50			

Observations provided at intervals of 1 min, for each of the north-looking and south-looking instrument yaws. O_3 = ozone, **u** = wind direction and speed.

the 'truth' (T) and of the assimilation model are similar, it is difficult to estimate their representativeness errors, and therefore these errors have been omitted. The implication of this is discussed in section 5(b).

(c) SWIFT observation characteristics

The viewing geometry of the SWIFT instrument means that the observation locations are dependent upon the yaw of the instrument on board the satellite (Lahoz *et al.* 2003). For each of the two months of the study, observations corresponding to both the north- and south-looking yaw states were produced. During the north-facing yaw, observations are available between approximately 87°N and 53°S, and during the south-facing yaw, observations are available between approximately 53°N and 87°S. For both yaws, the latitudinal spacing of the observations is about 3.5° near the equator and the longitudinal spacing varies from about 0.8° at 87°N or 87°S to about 5.5° at 53°N or 53°S.

The characteristics of the SWIFT wind and ozone measurements are discussed in the SWIFT Mission Requirements Document (MRD) (see, for example, Lahoz *et al.* (2003)). Based on theoretical considerations and modelling studies using the Canadian Middle Atmosphere Model (CMAM; Beagley *et al.* 1997), target values (that is those necessary to meet the science goals of SWIFT fully) and threshold values (those acceptable in order to make incremental improvements in our scientific understanding) have been derived for the SWIFT errors and resolution. Target values are more stringent than threshold values. As discussed in the SWIFT MRD, studies with the SWIFT instrument, radiative-transfer models, and retrieval models show that the contribution of SWIFT measurement noise to retrieval errors for wind and ozone in the altitude range 20–45 km is expected to be below target accuracy requirements. To make the results of this OSSE more robust, the error and resolution values (see below and Table 3) used in this experiment are equal to, or slightly more pessimistic than, the threshold values.

The wind and ozone measurements are represented by layer-mean values, with the layer centres located between 103 hPa and 0.8 hPa (approximately 16–50 km, at 2 km intervals). Ozone data are not available at and above 1.31 hPa. The averaging kernels calculated for SWIFT give an effective vertical resolution, comparable to that used in the OSSE (2 km).

The SWIFT ozone data were assumed to have a zero mean error, and an error STD assumed to be equivalent to 10% of a global-mean ozone value between 1.69 hPa and 55.0 hPa, 14% at 76.0 hPa and 18% at 103.0 hPa.

SWIFT wind errors were assumed to be 20 m s^{-1} at 0.8 hPa (50 km), linearly interpolated between 5 m s^{-1} and 20 m s^{-1} between 0.8 hPa and 2.9 hPa (50–40 km), 5 m s^{-1} between 2.9 hPa and 26 hPa (40–25 km), linearly interpolated between 5 m s^{-1} and 20 m s^{-1} between 26 hPa and 121 hPa (25–15 km), and 20 m s^{-1} at 121 hPa (15 km). These errors were used for both the zonal and meridional components.

The above SWIFT errors are the random component. The variation of SWIFT random errors with height is influenced by at least two factors: (i) the decrease of the signal-to-noise ratio with increasing height, and (ii) the higher line-of-sight errors in the upper and lower stratosphere. It is assumed that SWIFT has no systematic errors (or that, if they are known, they have been removed). Sources of known systematic errors include absorption-line characteristics. Once SWIFT is launched, a key part of the evaluation process will be the estimation of the SWIFT errors, including systematic errors. Following standard practice, the SWIFT errors are assumed to be uncorrelated, i.e. the error covariance matrix is diagonal.

In general, the SWIFT wind errors used in this study (see above) are slightly lower than those quoted for HRDI by Ortland *et al.* (1996). The random errors in the HRDI retrievals of zonal and meridional-wind components are typically 8 m s^{-1} in the mid-stratosphere (more in the tropics). However, it should be noted that the HRDI retrieval procedure includes a sequential estimation step that reduces the random errors by smoothing the winds between adjacent profiles.

(*d*) *The assimilation scheme*

The assimilation system used in the OSSE is based on the troposphere/stratosphere version of the operational Met Office three-dimensional variational data assimilation (3D-Var) scheme (Lorenc *et al.* 2000; Swinbank *et al.* 2002). The forecast model, known as the Unified Model (UM), has 40 levels from the surface to 0.1 hPa and has a horizontal resolution of $2.5° \times 3.75°$. In addition to the operationally assimilated variables, the assimilation system used here also includes ozone. The ozone assimilation scheme has produced satisfactory results when tested with HIRS channel 9 (Jackson and Saunders 2002) and SBUV* data.

The UM was modified to include a parametrization of photochemical sources and sinks of stratospheric ozone (Cariolle and Déqué 1986). Technical problems prevented the inclusion of a parametrization of ozone loss due to heterogeneous chemistry. In the context of this study, this omission is not particularly important since most heterogenous ozone destruction in the 1999/2000 NH winter took place in February and March, and very little in January and April (the OSSE trial periods).

Prior to insertion into the 3D-Var scheme, all observations are quality controlled by the Met Office observation processing system (OPS). Since many of the instruments shown in Tables 1 to 3 are future missions, and are thus not included in the operational OPS code, the code had to be modified so that observations from these instruments could be processed.

The 3D-Var code was extended to include the assimilation of retrieved profiles of ozone and RH, which was not present in the operational code. An investigation was made into the best way to represent the observation operator for these profiles. This operator is used to estimate observations given the model state, and is used in the calculation of the cost function, which is then minimized to produce an analysis.

The ozone and RH profiles are retrieved as mean values over observation layers. Following trial assimilation experiments, it was decided to use an observation operator that carries out linear interpolation of model profiles to the centres of observation layers.

(*e*) *Evaluation of the OSSE*

During assimilation experiments one needs to monitor whether the assimilation system is consistent with the Best Linear Unbiased Estimate (BLUE). A diagnostic
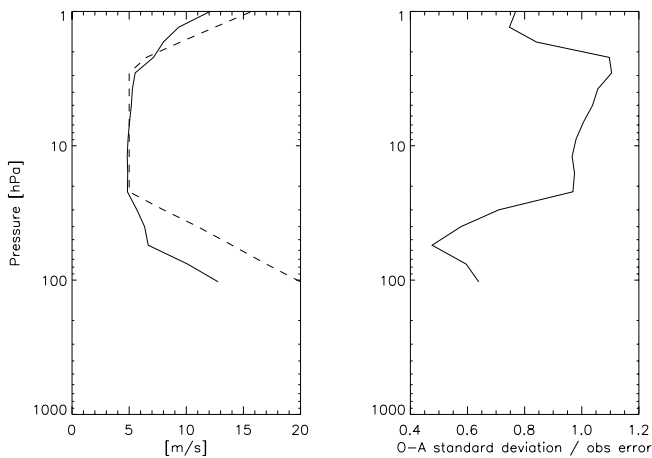
* Solar Backscattered Ultra Violet.

Figure 1. Profiles of (a) the standard deviation (STD) of the difference between SWIFT zonal-wind observations and collocated analysed zonal-wind values, O-minus-A (solid line) and the STD of the SWIFT zonal-wind component observation error (dashed line), and (b) the ratio of the STD of O-minus-A and the STD of the SWIFT observation error of the zonal-wind component. The results are for global coverage during January 2000 (m s^{-1}).

tool for testing whether analyses are consistent with a BLUE involves the D-minus-A difference (where D is the 'data' vector, comprised of observation, O, and background, B, components, and A is the analysis). If the analyses are consistent with a BLUE: (i) the bias of D-minus-A should be zero, (ii) the STD of D-minus-A should be less than or equal to the data error, and (iii) as the analyses approach the 'truth' (T), the STD of D-minus-A should approach the data error from below (Struthers *et al.* 2002; Talagrand 2003).

Often, for convenience, instead of the STD of D-minus-A, the bias and STD of O-minus-A are computed. In this case, the requirements for consistency with a BLUE are that the bias of O-minus-A be zero (or 'small' in comparison with the observation errors), and that the STD of O-minus-A be less than the observation error STD.

Figure 1 shows a sample result of a BLUE test for the SWIFT zonal-wind component using the SW configuration. Many similar results have been presented by Lahoz *et al.* (2003), confirming that the SWIFT zonal-wind analyses are consistent with a BLUE throughout the stratosphere and for all latitudinal ranges considered (e.g. global, tropics, mid and polar latitudes).

It is often desirable (and commonly assumed) that the errors in the observations and background are normally distributed, unbiased and mutually uncorrelated, but it is very difficult to check this directly. However, a necessary condition is that their difference O-minus-B is normally distributed. A histogram of these differences can test whether the differences are (approximately) normally distributed and unbiased. Typically, the STD of the O-minus-B histograms will be larger than that of the O-minus-A histograms. This is because the analysis errors are smaller (on average) than the background errors.

An investigation of histograms shows that the O-minus-B differences for SWIFT zonal winds are normally distributed throughout the stratosphere (Lahoz *et al.* 2003). These results support the assumption that the errors in O and B are normally distributed and uncorrelated. These results also show that the histograms for O-minus-A have smaller STDs than those for O-minus-B, suggesting that the analyses are behaving in a reasonable manner.
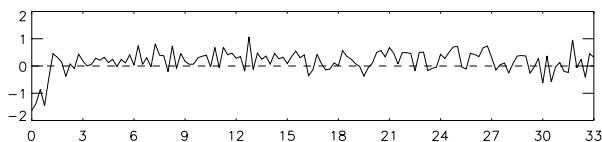
Figure 2.   Time-series of the difference between the SWIFT zonal-wind observations and collocated analysed zonal-wind values, O-minus-A. The results are for global coverage at 12 hPa starting at 18 UTC 1 January 2000. The *x*-axis is the time (days), the *y*-axis is wind velocity (m s^{-1}), and the dashed line indicates the zero O-minus-A difference.

Figure 2 shows a sample month-long time series of O-minus-A differences for the SWIFT zonal-wind component using the SW configuration. Combining such information with the time-series shown by Lahoz *et al.* (2003), it is possible to infer that, overall, the assimilation of SWIFT winds produces satisfactory analyses throughout most of the stratosphere. Biases increase in the upper stratosphere because differences between the Met Office and the ECMWF (used to derive the 'truth', T) background-wind fields increase, and because the SWIFT errors are larger. The time-series also suggest a spin-up period of one to two days in the stratosphere for the SWIFT zonal-wind component. The small O-minus-A differences observed after the spin-up period in the lower/mid-stratosphere suggest a stable assimilation system.

The results of comparable investigations of the SWIFT meridional wind are very similar to those for SWIFT zonal wind. Analogous results for SWIFT ozone suggest that it is consistent with a BLUE in the mid and lower stratosphere (100–10 hPa), but is not in the upper stratosphere (10–1 hPa). This is because there is a bias in the ozone background in this region and also because the ozone background-error covariance here is too small. A sensitivity test, in which the stratospheric ozone background-error covariance was increased, produced ozone analyses that were closer to a BLUE in the upper stratosphere.

3.   IMPACT OF SWIFT MEASUREMENTS

In this section, the impact of SWIFT measurements is assessed by computing statistics of the differences between the SW and NS analyses and the 'truth', T, for zonal and meridional-wind components, and ozone. We focus on the zonal wind and briefly discuss the impact for meridional-wind and ozone. Further details can be found in the paper by Lahoz *et al.* (2003).

To provide a comprehensive measure of the impact of SWIFT data, a number of investigations were carried out. First, we examined histograms and monthly mean fields for the SW-minus-T and NS-minus-T differences, together with monthly-mean fields for T, NS and SW. Such an examination for winds suggests that the SW and NS analyses are realistic, and that the SWIFT wind measurements tend to have most positive impact on the analyses in the tropics, and some impact in the extratropics (especially at high latitudes in wintertime). The impact decreases with decreasing height, being large at 1 hPa and 10 hPa and very small at 100 hPa.

The SW and NS ozone analyses are also realistic. The histograms and monthly means show few differences between the two analyses, but there are indications of positive impacts in the wintertime extratropics and the tropics, especially at 10 hPa and 100 hPa.

The histograms and monthly means mainly provide qualitative information on the impact of SWIFT. To provide a more quantitative evaluation, root-mean-square

(RMS) values for the SW-minus-T and NS-minus-T differences were calculated, and significance tests were carried out on these differences. The RMS and significance tests also provide information on how the differences vary spatially, temporally and with SWIFT look. For this reason, we devote the rest of section 3 to these two tests.

(a)   RMS and significance tests

The RMS tests were done at a series of pressure levels (100 hPa, 50 hPa, 10 hPa, and 1 hPa). The study focused on the variation of the ratio $\rho = \text{RMS(SW-minus-T)}/\text{RMS(NS-minus-T)}$ over different latitudinal ranges and different SWIFT looks. A value of ρ less than one indicates a beneficial impact from SWIFT (although this may not be significant in the statistical sense).

The significance tests were performed on the 2–31 January and 1–29 April SW-minus-T and NS-minus-T differences at 100 hPa, 50 hPa, 10 hPa and 1 hPa. In principle, the datasets have sufficient data to use the normal distribution to test the null hypothesis that the calculated means are equal, given the STD of the datasets. It is assumed that the normal distribution holds for the null hypothesis.

To provide robustness to the study, the following tests were carried out: (i) two-sided significance tests at the 0.95 confidence limit (CL) assuming that the members within each dataset (SW-minus-T; NS-minus-T) are independent and that the datasets are independent (i.e. there is no correlation between the datasets), (ii) two-sided significance tests at the 0.95 CL, including the effects of persistence and correlation between the datasets (see Wilks (1995), chapter 5), and (iii) the CL at which the null hypothesis would be rejected in a two-sided significance test (including effects of persistence and correlation between the datasets). In this paper we show only results from test (ii), although it should be noted that the results are not sensitive to the treatment of the datasets.

We can discount in these results the effect from the bias between the ECMWF and Met Office systems, because this bias is removed when comparing the SW-minus-T and NS-minus-T differences.

(b)   Impact on zonal-wind analyses

Table 4 shows the RMS statistics for the zonal wind for January (including both SWIFT looks). This table and the results presented by Lahoz *et al.* (2003), suggest that the SWIFT wind measurements are most likely to have an impact on the zonal-wind analyses in the tropics in the mid and upper stratosphere (50–1 hPa).

The impact of SWIFT on the tropical zonal-wind analyses is also illustrated in Fig. 3, which shows the results of the significance test for zonal wind for January at 50 hPa, 10 hPa and 1 hPa. This figure, and the results presented by Lahoz *et al.* (2003) show that the SWIFT wind measurements have a significant positive impact (i.e. the shaded region in Fig. 3) on the zonal-wind analyses at the 0.95 CL in the tropics at 50 hPa, 10 hPa and 1 hPa. The region of significant impact extends over a much larger area when the CL is reduced to 0.75. In addition, (for 50 hPa and 1 hPa only) the regions of significance at the 0.99 CL tend to be comparable with those at the 0.95 CL. However, at 100 hPa (not shown) SWIFT wind measurements are unlikely to have a significant impact on the zonal wind at the 0.95 CL.

The regions where the NS analyses are closer to T than the SW analyses are very small in area and are generally not significant at the 0.95 CL.

The large impact of SWIFT measurements on the tropical mid- to upper-stratospheric zonal-wind analyses is related to the fact that, in this region, there is less direct
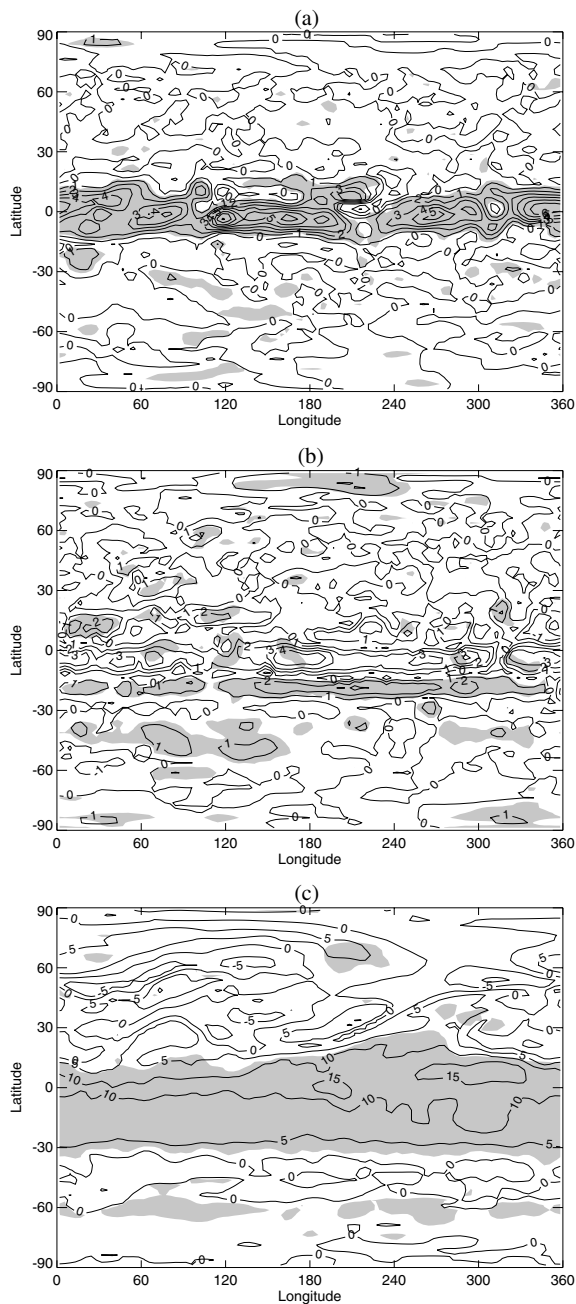
Figure 3. (a) Plot of the difference $|\text{NS-minus-T}| - |\text{SW-minus-T}|$ for the monthly-mean zonal wind (m s^{-1}) at 50 hPa for 2–31 January. The shading indicates (i) where both the difference between the monthly means for NS-minus-T and SW-minus-T is significant at the 0.95 confidence level, and (ii) where the SW analyses are closer to the T values than the NS analyses. (b) As (a), but at 10 hPa. (c) As (a), but at 1 hPa. See text for further explanation.

TABLE 4. THE ρ STATISTIC (SEE TEXT) FOR THE ZONAL WIND IN JANUARY: (a) THE NORTH-
ERN AND SOUTHERN HEMISPHERE LOOKS, (b) THE NORTHERN HEMISPHERE LOOK (SWIFT
MEASUREMENTS FROM 53°S TO 87°N), AND (c) THE SOUTHERN HEMISPHERE LOOK (SWIFT
MEASUREMENTS FROM 87°S TO 53°N)

	Global	90°S–60°S	60°S–30°S	30°S–30°N	30°N–60°N	60°N–90°N
(a) 2–31 January						
100 hPa	0.95	0.91	0.92	0.95	0.94	0.96
50 hPa	0.61	0.95	0.96	0.55	0.92	0.90
10 hPa	0.66	0.90	0.76	0.52	0.86	0.82
1 hPa	0.66	0.83	0.91	0.49	0.99	0.86
(b) 2–17 January						
100 hPa	0.94	1.03	0.95	0.92	0.97	1.04
50 hPa	0.67	0.95	0.99	0.62	0.91	1.00
10 hPa	0.55	1.00	0.87	0.39	0.78	0.78
1 hPa	0.64	0.90	0.91	0.58	0.86	0.87
(c) 18–31 January						
100 hPa	0.94	0.81	0.93	0.96	0.92	1.00
50 hPa	0.63	0.98	1.00	0.53	0.93	1.02
10 hPa	0.76	0.76	0.69	0.57	0.94	0.98
1 hPa	0.71	0.95	0.92	0.47	0.89	0.82

wind information from sondes than, for example, in the NH mid-latitudes, and none at all above 10 hPa. Furthermore, in the tropics, geostrophy breaks down and no indirect wind information from temperature measurements aboard nadir sounders is available. Conversely, in the extratropics this indirect information exists and, in addition, there is more direct wind information from sondes—hence the impact of SWIFT on the analyses is smaller.

Table 4 also shows that SWIFT wind measurements may have an impact on the zonal-wind analyses in the mid and upper stratosphere extratropics when SWIFT data is available (this depends on the SWIFT look) and for autumn/winter/spring conditions (typically associated with flow regimes that are changing relatively fast). Consistent with this, Fig. 3 shows that, in the extratropics (at 10 hPa and 1 hPa), there are patches of significance in autumn/winter mid-latitudes and polar latitudes. However, the significance may be less robust than in the tropics.

The above result suggests that SWIFT wind measurements are needed to capture these fast-changing flow regimes. Conversely, this suggests that indirect wind information from temperature measurements aboard nadir sounders may not be enough to capture these flow regimes.

(c)   *Impact on meridional-wind and ozone analyses*

RMS errors show that the impact of SWIFT wind measurements on meridional-wind analyses is smaller than that for zonal wind. In addition, unlike for zonal wind, the impact on the meridional-wind analyses is not clearly larger in the tropics than elsewhere. SWIFT wind measurements may have a significant impact on meridional wind at 0.95 CL at 10 hPa in some regions but, at 100 hPa, 50 hPa and 1 hPa, the regions of significance at the 0.95 CL are small and very isolated. At all the pressure levels considered, the SWIFT wind measurements have a significant impact on the meridional wind over a reasonably wide area at the 0.75 CL, but the spatial distribution of these areas is patchy.

The small positive impact of SWIFT measurements on the meridional-wind analyses may be partly due to the meridional winds having a smaller magnitude, while the

observation error applied to both zonal and meridional wind components is the same. This leads to smaller analysis increments for the meridional wind component.

Another possible explanation is that, in the Met Office 3D-Var scheme, the background wind errors are treated as a combination of errors in the non-divergent wind field and errors in the irrotational wind field. Since the non-divergent wind (corresponding to the gradient of the stream function) tends to be predominantly east–west, one might speculate that analysis increments to the non-divergent wind would tend to be predominantly reflected in the zonal-wind component. Since the irrotational part of the wind and the corresponding background error are generally smaller, the analysis increments will also tend to be smaller.

The RMS errors for the ozone analyses show that the SWIFT ozone measurements tend to have an impact in stratospheric regions where the vertical gradient of ozone is relatively high, namely at 10 hPa in the tropics and extratropics (near the ozone peak), and at 100 hPa in the tropics (near the tropopause). Consistent with this, the significance tests show that SWIFT ozone measurements may have a significant impact on the ozone analyses at the 0.95 CL at 10 hPa in the tropics and isolated regions in the extratropics, and at 100 hPa in the tropics (April only). Overall, the impact at 100 hPa and 10 hPa appears to be more significant than at 50 hPa or 1 hPa. This suggests that the higher vertical resolution of SWIFT can better resolve the higher vertical gradients of ozone.

The RMS errors also suggest that the SWIFT ozone measurements tend to have an impact at 50 hPa in the extratropics for autumn/winter/spring conditions. This is corroborated by the significance tests, which show significant impacts in mid-latitudes at 50 hPa, and also in mid-latitudes and autumn/winter/spring polar latitudes at 100 hPa. Since the ozone field is chiefly determined by transport at these levels, it is possible that the improved zonal-wind analyses under these conditions can contribute to the improvement in the ozone analyses.

4.   A COMPARISON OF WIND ANALYSES

In this section we illustrate further the potential benefits of SWIFT observations by investigating tropical winds (section 4(a)) and zonal-wind wintertime variability (section 4(b)).

(a)   *Tropical winds*

Figure 4 shows the latitude/pressure monthly means for the zonal wind for January for T, SW-minus-T and NS-minus-T. The SW and NS zonal-wind analyses are generally much closer to each other than to T. The most marked differences between the two analyses are seen in the tropical mid and upper stratosphere, where the SWIFT errors are relatively low and SWIFT winds have more of an impact (see section 2(c) and Table 4). There, the NS-minus-T difference is about three times larger than the SW-minus-T difference. In April the SW-minus-T difference is also much smaller than the NS-minus-T difference in the tropical upper stratosphere; however, these differences are of opposite sign to those in January. This easterly bias is explained by the fact that the UM fails to simulate the westerly phase of the semi-annual oscillation (SAO) in April—instead, it only simulates weak easterlies (Amodei *et al.* 2001).

Note that Fig. 4(b) shows small differences for SW-minus-T in the neighbourhood of 10 hPa. This is consistent with the results in Fig. 2, which show small differences for O-minus-A (SW analyses) at 12 hPa in January.

We also note that both NS and SW analyses produced a realistic simulation of the Brewer–Dobson circulation. However, the circulations calculated from both analyses
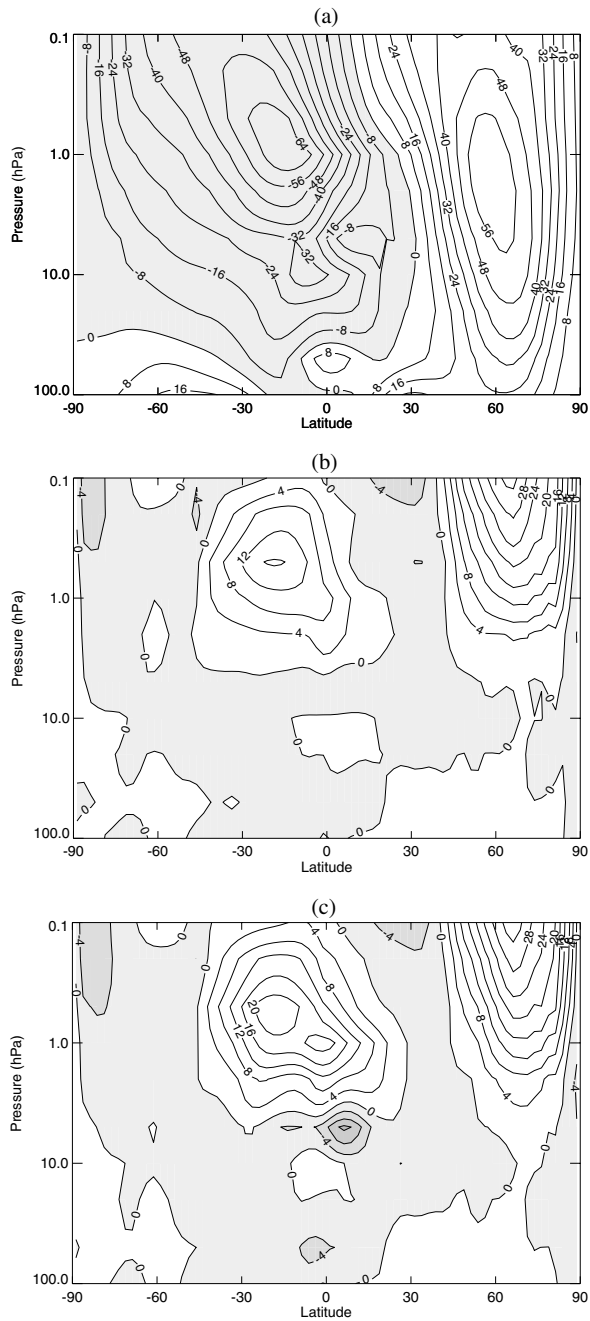
Figure 4. Latitudinal cross-section against pressure (hPa) of the monthly-mean zonal wind (m s^{-1}) during January 2000 for (a) the 'truth' (T) (the shading indicates easterly winds), (b) the difference SW-minus-T (the shading indicates that the SW analyses are more easterly than T), and (c) the difference NS-minus-T (the shading indicates that the NS analyses are more easterly than T). See text for further explanation.

are very similar to each other largely because, for the meridional wind, the SW-minus-T and NS-minus-T differences are broadly very similar to each other in both January and April.

Figure 5 shows the latitude/day time-series for the zonal wind for January at 10 hPa for T, SW-minus-T and NS-minus-T. It can be seen that, in the tropics, the SW analyses are closer to T than the NS analyses, with the largest discrepancy between the NS-minus-T and SW-minus-T differences occurring at the beginning of the experiment for both January and April (not shown). The clear reduction in the NS-minus-T differences in the first few days of the experiment stems from the large differences between the initial conditions, which are due to difficulties in modelling the QBO and the paucity of tropical wind observations. It is reasonable to assume that there may be a large error in these initial conditions, and thus that the synthetic sonde observations of tropical wind near 10 hPa used in the OSSE will differ greatly from the actual observations used to generate the initial conditions. The impact is to drag the NS analyses toward T (as seen in Fig. 5).

This behaviour suggests a timescale of about 10 days in the NS configuration for direct wind information to propagate from the isolated tropical sondes to the whole tropics. The result is consistent with previous studies that examined the impact of tropical sondes in the Met Office stratospheric data assimilation system. A similar reduction in NS-minus-T differences is not seen at other levels because in the lower stratosphere the difference between NS and T initial conditions is much smaller, and because above the 10 hPa level no sonde wind observations are available.

After 10 days, although the NS analyses markedly improve in the tropics, they are still further away from T than the SW zonal-wind analyses. This suggests that, at 10 hPa, neither the sondes nor the nadir temperature sounders can provide as much information as the SWIFT wind measurements.

(b)   Wintertime variability

To assess the wintertime variability in the wind for T and the analyses, we calculated the STD of the monthly mean for the zonally averaged zonal wind.

Figure 6 shows the variability for the zonal component of wind for January. This figure and the results presented by Lahoz *et al.* (2003) show that wintertime variability is largest in the region of the polar-night jet (PNJ). The pattern of variability is similar for T and the SW and NS analyses, but the analyses show smaller variability than T. This may be associated with the analyses overestimating the strength of the PNJ with respect to T. A likely reason for this overestimate is the difference between the atmospheric representations of the Met Office and ECMWF analyses (see section 2).

The PNJ region is where the flow regime is most likely to change. Following on from the results shown in section 3, the analyses are more likely to represent this variability better when SWIFT observations are available. Figure 6 shows that this is indeed the case; in the PNJ region, the variability in SW is closer to that of T (by up to a factor of three) than that in NS.

In summary, these results demonstrate that the SWIFT wind observations will tend to improve the representation of wintertime variability in the extratropical stratosphere. This, and the improvement in tropical wind information shown in section 4(a) is likely to be more important for NWP than for improving the tropical wind climatology. Although one of the goals of SWIFT is to improve tropical wind climatologies, this OSSE cannot address this issue directly as the time periods analysed are too short.
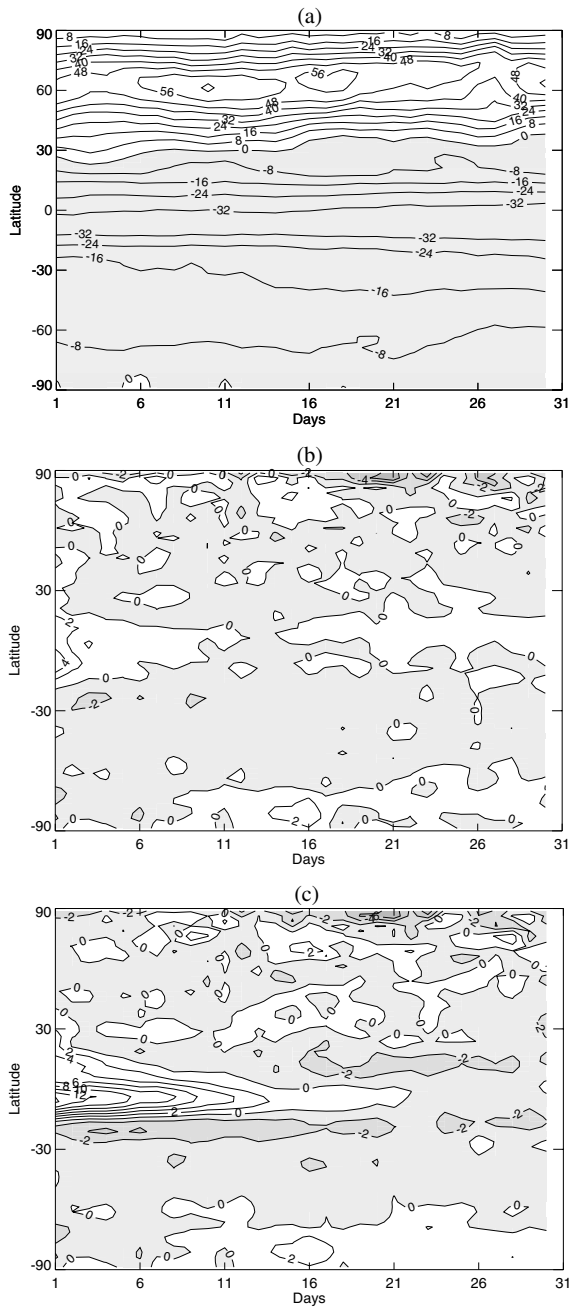
Figure 5.   Time-series of the latitudinal variation of the zonal-wind (m s^{-1}) during January 2000 at 10 hPa for (a) the 'truth' (T) (the shading indicates easterly winds), (b) the difference SW-minus-T (the shading indicates that the SW analyses are more easterly than T), and (c) the difference NS-minus-T (the shading indicates that the NS analyses are more easterly than T). See text for further explanation.
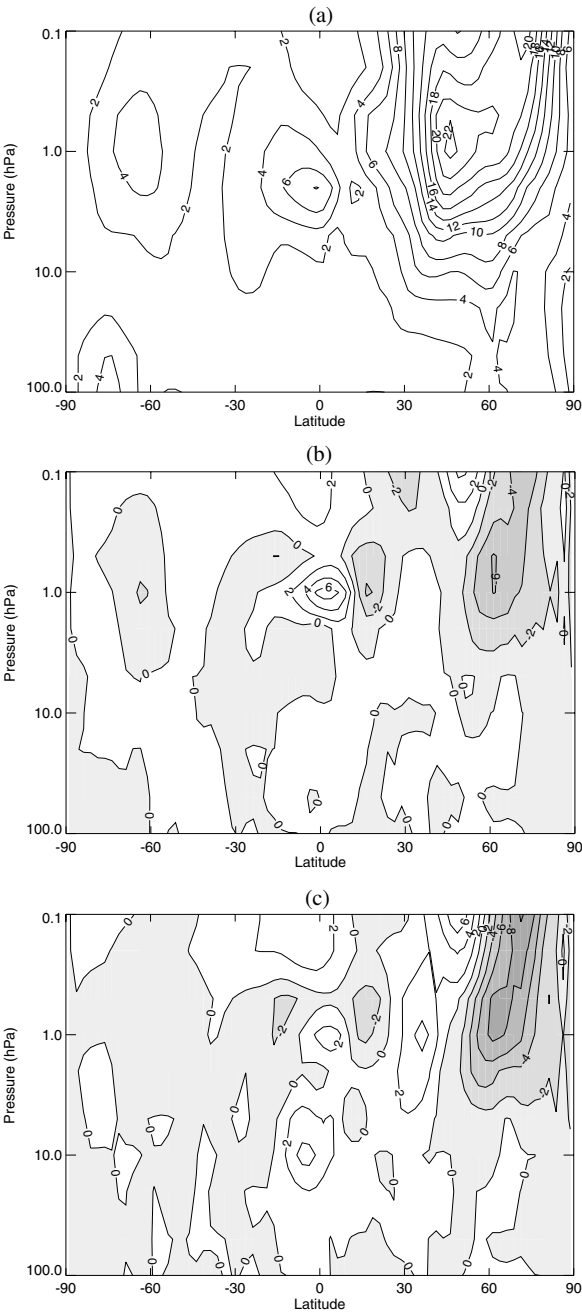
Figure 6. Latitudinal cross-section against pressure (hPa) of the zonal-wind variability (m s^{-1}) during January 2000 for (a) the 'truth' (T), (b) the difference SW-minus-T (the shading indicates that the SW analyses are more easterly than T), and (c) the difference NS-minus-T (the shading indicates that the NS analyses are more easterly than T). See text for further explanation.

5. DISCUSSION ON THE MERITS OF SWIFT

(*a*) *Impact and scientific benefits of SWIFT*

The results presented earlier in this paper demonstrate a positive benefit from the assimilation of simulated SWIFT wind observations. The largest impact of SWIFT observations on zonal-wind analyses is seen in the tropical middle and upper stratosphere. There is also an impact in the extratropical upper stratosphere in cases where SWIFT data are available and the flow regime is changing relatively fast. In the absence of SWIFT wind observations, winds need to be inferred from temperature soundings via the thermal-wind relation, but clearly in these situations this relation is not accurate. A previous study (Boorman *et al.* 2000) found only marginal benefit from assimilating winds measured by HRDI, but they were substantially less accurate than the SWIFT measurements are expected to be (see section 2(*c*)). The results shown here indicate that the SWIFT observations will be accurate enough to have a positive impact.

Given the impact on the zonal-wind analyses, it is perhaps surprising that there was little positive impact of SWIFT observations on the meridional-wind analyses. Possible explanations are a larger relative error for the meridional-wind observations, and the predominance of non-divergent analysis increments, which tend to be along the zonal-wind component. The irrotational part of the wind, and the corresponding background error, is generally smaller and the analyses increments also tend to be smaller. So, SWIFT wind observations may be more effective at correcting the zonal-wind component. Another possible explanation is that, since the meridional-wind component is usually smaller than the zonal component, the relative impact of wind observations on the meridional wind is larger (and perhaps non-optimal).

There is also some indication that SWIFT observations may improve the representation of the Brewer–Dobson circulation, but this is tentative, reflecting the small impact of SWIFT data on the meridional-wind analyses.

Turning to ozone, simulated SWIFT observations have an impact on ozone analyses at stratospheric levels with relatively high vertical gradients of ozone, i.e. near 10 hPa and 100 hPa. This shows the benefit of the relatively high vertical resolution of the SWIFT ozone measurements. There are also indications of further benefits of SWIFT observations in regions where the flow regime is changing quickly; this probably reflects improved modelling of ozone transport associated with the presence of SWIFT observations. Note, however, that such events are patchy and not always statistically significant.

In summary, the results of this study suggest that SWIFT data will improve the quality of stratospheric analyses, particularly the tropical winds. Improved climatologies could then be used to evaluate GCMs in the middle atmosphere better. This OSSE cannot directly address the issue of climatologies, as the time periods are too short, but it is clear from the results here that there is likely to be a benefit for NWP in the stratosphere. SWIFT data are also likely to benefit transport studies in which trajectories are calculated from regular analyses.

(*b*) *Caveats of using a 'reduced OSSE'*

It has already been mentioned in section 1 that, because of limited resources, we decided to use an approach referred to as a 'reduced OSSE'. Although we have judged that this approach is acceptable for evaluating an innovative instrument such as SWIFT, there are a number of caveats that merit further discussion.

First, we simulated retrieved profiles of temperature instead of simulating satellite radiances. The assimilation of radiances, rather than retrieved profiles, has been shown

to lead to improvements in analyses and forecast skill (e.g. English *et al.* 2000). The chief advantages of using radiances are the avoidance of inappropriate a-priori information (which is inherent in most retrieved profiles) and better representation of observation errors. The former is not an issue with this OSSE, because the retrieved profiles were calculated directly from the reference atmosphere. However, the representation of the observation-error covariances is rather crude, and might erroneously alter the impact of satellite soundings on the analyses.

When SWIFT is launched, one can anticipate that an operational system would assimilate radiances from both satellite instruments such as AMSU-A and IASI, and retrieved wind profiles from SWIFT. The use of radiances, rather than the retrievals used in this OSSE, is unlikely to have an impact on the wind analyses in the tropics, since the thermal-wind relation does not hold there. However, extratropical wind analyses are likely to be affected by the assimilation of radiances. But, since it is suggested that SWIFT wind observations have a positive impact when the thermal-wind relation is inaccurate (e.g. in fast-changing flow regimes), this positive impact is likely to remain.

Second, the experiments reported here have been run at a rather low horizontal resolution, T42 for the reference atmosphere, and $2.5° \times 3.75°$ for the Met Office stratospheric model. As stated in section 2(b), only representativeness errors for aircraft, sonde and surface data were included. These types of observations are point-like measurements and thus their representativeness errors are likely to be large relative to instrumental and background errors (Cohn 1997; Sparling *et al.* 1998). On the other hand, representativeness errors for the other observations used in the OSSE, which are from satellite instruments, are likely to be smaller. In particular, the omission of representativeness errors for SWIFT is likely to have only a small impact on the conclusions of this study. In any event, given that all the observations are simulated from a reference atmosphere that has a similar resolution to that of the assimilation model, the importance of representativeness error will be minor here. Clearly, though, if a higher-resolution reference atmosphere had been available, it would have been possible, and advisable, to calculate realistic representativeness errors for all observation types.

By the time that SWIFT is launched, the horizontal resolution of the model used at the Met Office for stratospheric data assimilation is likely to be much higher. One could have anticipated this by running the OSSE at a higher resolution; however, this would have entailed extra expense—both to set up a non-standard higher-resolution assimilation system, and also to run the assimilation experiments. With a higher-resolution model, one would decrease the thinning of satellite data, which could greatly increase the amount of AMSU-A and IASI data assimilated. In turn, these extra data would impact the stratospheric wind analyses outside the tropics, and may change the degree of benefit obtained from the SWIFT data. However, the impact of SWIFT on tropical-wind analyses is likely to remain robust.

The effect of improved horizontal resolution on the ozone analyses is likely to be small, since that impact is chiefly explained by the greater vertical resolution of the SWIFT ozone measurements.

6.   CONCLUSIONS AND RECOMMENDATIONS

The results shown in this paper strongly suggest that SWIFT observations would positively impact zonal-wind analyses in the tropical stratosphere, and in the extratropical upper stratosphere when the flow regime is changing relatively fast and SWIFT observations are available. There is also a strong suggestion that they will positively impact stratospheric ozone analyses in regions where the vertical gradient is relatively

high. Such impacts are likely to provide benefits to the wider scientific community. The various results presented in sections 3 and 4 are consistent with each other and with the conclusions above. Furthermore, most of the conclusions remain robust, even after consideration of the caveats of the 'reduced OSSE' approach (section 5(b)). Accordingly, we strongly recommend that the development, construction and subsequent launch of the SWIFT instrument be implemented.

Despite the robustness of the results presented here, it should be stressed that the SWIFT observations under examination were innovative, being a completely different observation type from the other observations available in the operational observing system. However, if one were evaluating a new observation type that embodied more subtle improvements over existing observations, then the 'reduced OSSE' approach would be likely to be inadequate. Instead, one would need to construct a full OSSE to obtain credible results. Such a task is very expensive, perhaps requiring an order of magnitude more resources than were used for this study. It is clear that OSSEs can play a very important role in the assessment of proposed space missions. Thus we recommend that space agencies (possibly in conjunction with other meteorological centres) fund the development of a full OSSE capability which could then be used to assess a variety of future proposed space missions in a more credible and sophisticated manner.

REFERENCES

Amodei, M., Pawson, S., Scaife, A. A., Langematz, U., Lahoz, W., Li, D. M. and Simon, P.   2001   The SAO and Kelvin waves in the EuroGRIPS GCMs and the UK Met. Office analyses. *Ann. Geophys.*, **19,** 99–114

Atlas, R. M.   1997   'Observing system simulation experiments: Methodology, examples and limitations'. CGC/WMO Workshop, Geneva 7–9 April 1997, WMO TD No. 868., World Meteorological Organization, Geneva, Switzerland

Beagley, S. R., de Grandpré, J., Koshyk, J. N., McFarlane, N. A. and Shepherd, T. G.,   1997   Radiative–dynamical climatology of the first-generation Canadian middle atmosphere model. *Atmos.–Ocean*, **35,** 293–331

Boorman, P. M., Swinbank, R. and Ortland, D. A.   2000   'Assimilation of directly measured stratospheric winds into the Unified Model'. Forecasting Research Tech. Report 332, Met Office, Exeter, UK

Cariolle, D. and Déqué, M. J.   1986   Southern hemisphere medium-scale waves and total ozone disturbances in a spectral general circulation model. *J. Geophys. Res.*, **91,** 10825–10884

Chipperfield, M. P.   1999   Multi-annual simulations with a three-dimensional chemical transport model. *J. Geophys. Res.*, **104,** 1781–1805

Cohn, S. E.   1997   An introduction to estimation theory. *J. Meteorol. Soc. Japan.*, **75,** 257–288

Dobbie, J. S., Cann, M. W. P., Rochon, Y., Gault, B., Ward, W., Shepherd, G. G. and Rowlands, N.   1996   'Emission line selection for the SWIFT instrument'. Canadian Aeronautics and Space Institute conference, Ottawa, November 1996

EC   2000   'Severe stratospheric ozone depletion in the Arctic'. European Commission DGXII Press Release 5 April 2000 (available from http://www.nilu.no/projects/theseo2000)

English, S. J., Renshaw, R. J., Dibben, P. C., Smith, A. J., Rayer,P. J, Poulsen, C., Saunders, F. W. and Eyre, J. R. | 2000 | A comparison of the impact of TOVS and ATOVS satellite sounding data on the accuracy of numerical weather forecasts. *Q. J. R. Meteorol. Soc.*, **126,** 2911–2931

Hays, P. B., Abreu, V. J., Dobbs, M. E., Gell, D. A., Grassl, H. J. and Skinner, W. R. | 1993 | The High-Resolution Doppler Imager on the Upper Atmosphere Research Satellite. *J. Geophys. Res.*, **98,** 10713–10723

Jackson, D. R. and Saunders, R. | 2002 | 'Ozone data assimilation: Preliminary system'. Forecast Research Tech. Report 394, Met Office, Exeter, UK

Lahoz, W. A., Brugge, R., Migliorini, S., Lary, D., Lee, A., Swinbank, R. and Jackson, D. | 2003 | 'SWIFT—Application of stratospheric wind measurements to process studies and meteorology'. Final Report, March 2003, ESTEC Contract No. 15344/01/NL/MM. Available from ESA Publications Division, PO Box 299, 2200 AG Noordwijk, The Netherlands

Lorenc, A. C., Ballard, S. P., Bell, R. S., Ingleby, N. B., Andrews, P. L. F., Barker, D. M., Bray, J. R., Clayton, A. M., Dalby, T. D., Li, D., Payne, T. J. and Saunders, F. W. | 2000 | The Met Office global three-dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc.*, **126,** 2991–3012

Ortland, D. A., Skinner, W. R., Hays, P. B., Burrage, M. D., Lieberman, R. S., Marshall, A. R. and Gell, D. A. | 1996 | Measurements of stratospheric winds by the High Resolution Doppler Imager, *J. Geophys. Res.*, **101,** 10351–10363

Randel, W. J., Fleming, E., Geller, M., Gelman, M., Hamilton, K., Karoly, D., Ortland, D., Pawson, S., Swinbank, R., Udelhofen, P., Baldwin, M., Chanin, M.-L., Keckhut, P., Simmons, A. and Wu, D. | 2002 | 'SPARC intercomparison of middle atmosphere climatologies'. WCRP, WMO/TD, SPARC Report 3, September 2002, World Meteorological Organization, Geneva, Switzerland

Shepherd, G. G., McDade, I. C. and Gault, W. A. | 1998 | 'SWIFT Explorer: Stratospheric Wind Interferometer For Transport studies'. Proposal to ESTEC for an Earth Explorer Opportunity Mission, Proposal COP27, European Space Research and Technology Centre, Noordwijk, The Netherlands

Sparling, L. C., Douglass, A. R. and Schoeberl, M. R. | 1998 | An estimate of the effect of unresolved structure on modelled ozone loss from aircraft observations of ClO. *Geophys. Res. Lett.*, **25,** 305–308

Struthers, H., Brugge, R., Lahoz, W. A., O'Neill, A. and Swinbank, R. | 2002 | Assimilation of ozone profiles and total ozone column measurements into a global general circulation model. *J. Geophys. Res.*, **107,** doi: 10.1029/2001JD000957

Swinbank, R. and O'Neill, A. | 1994 | A stratosphere–troposphere data assimilation system. *Mon. Weather Rev.*, **122,** 686–702

Swinbank, R., Ingleby, N. B., Boorman, P. M. and Renshaw, R. J | 2002 | 'A 3D variational data assimilation system for the stratosphere and troposphere'. Forecasting Research Scientific Paper 71, Met Office, Exeter, UK

Talagrand, O. | 2003 | 'A posteriori validation of assimilation algorithms'. In *Data assimilation for the earth system*, Eds. R. Swinbank, V. Shutyaev and W. A. Lahoz. Kluwer, The Netherlands

Wilks, D. S. | 1995 | *Statistical methods in the atmospheric sciences.* Academic Press

WMO | 1999 | 'Scientific Assessment of Ozone Depletion', Global Ozone Research and Monitoring Programme Report 44, World Meteorological Organization, Geneva, Switzerland